RAW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY
SYSTEMS
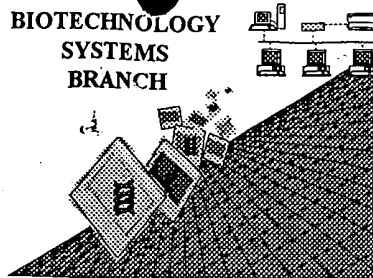BRANCH

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/841,157_

Source: _OIPE_

Date Processed by STIC: _10/11/2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

ERROR DETECTED   SUGGESTED CORRECTION   SERIAL NUMBER: 09/841,157

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWAR

1_____Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length

Sequence(s)_____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7_____Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10_____Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11_____Use of <220>

Sequence(s) 1-44 missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13_____Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

OIPE

```
        RAW SEQUENCE LISTING                DATE: 10/11/2001
        PATENT APPLICATION:  US/09/841,157  TIME: 10:40:08

        Input Set : A:\US09841157SEQLIST.txt
        Output Set: N:\CRF3\10112001\I841157.raw
```

```
        3 <110> APPLICANT: Taitec Co., Ltd.
W-->    4 <120> TITLE OF INVENTION: Method for identifying an organism by its genotype
C-->    5 <140> CURRENT APPLICATION NUMBER: US/09/841,157
C-->    5 <141> CURRENT FILING DATE: 2001-04-25
E-->    5 <160> NUMBER OF SEQ ID:
W-->    0 <130> FILE REFERENCE:
E-->    5 <160> NUMBER OF SEQ ID NOS:
```

ERRORED SEQUENCES

Does Not Comply
Corrected Diskette Needed

```
W-->  6 <210> SEQ ID NO: 1
      7 <211> LENGTH: 204
      8 <212> TYPE: DNA
W-->  9 <211> LENGTH: unknown
W--> 10 <213> ORGANISM:
E--> 10 <400> SEQUENCE: 0
     11 tgctacgtct cttccgatgc tgtctttcgc tgctgagggt gacgatcccg caaaagcggc 60
     12 ctttgactcc ctgcaagcct cagcgaccga atatatcggt tatgcgtggg cgatggttgt 120
     13 tgtcattgtc ggcgcaacta tcggtatcaa gctgtttaag aaattcacct cgaaagcaag 180
     14 ctgataaacc gatagaattc aagg 204
     55 <210> SEQ ID NO: 8
E--> 56 <211>12
     57 <212> TYPE: DNA
     58 <211> LENGTH: unknown
W--> 59 <213> ORGANISM:
W--> 59 <400> SEQUENCE: 8
E--> 60 aaatttaaat tt 12
     61 <210> SEQ ID NO: 9
     62 <211> LENGTH: 12
     63 <212> TYPE: DNA
W--> 64 <211> LENGTH: unknown
W--> 65 <213> ORGANISM:
W--> 65 <400> SEQUENCE: 9
E--> 66 aattaattaa tt
    187 <210> SEQ ID NO: 30
    188 <211> LENGTH: 12
    189 <212> TYPE: DNA
W--> 190 <211> LENGTH: unknown
W--> 191 <213> ORGANISM:
W--> 191 <400> SEQUENCE: 30
E--> 192 gatcacctcc tta 13)
    193 <210> SEQ ID NO: 31
    194 <211> LENGTH: 12
    195 <212> TYPE: DNA
W--> 196 <211> LENGTH: unknown
W--> 197 <213> ORGANISM:
```

see following pages

many errors

Please consult Sequence Rules

SEQUENCE LISTING ~~delete~~

<110> Taitec Co., Ltd.

<120> Method for identifying an organism by its genotype

<160> ← mandatory response needed (44 shown) → <140> <141>

<210> 1

<211> 204

<212> DNA

<213> <211> unknown *see item 11 on Erra Summary sheet*

global error

insert these mandatory numeric identifiers and response

<400> 1

```
tgctacgtct cttccgatgc tgtctttcgc tgctgagggt gacgatcccg caaaagcggc 60
ctttgactcc ctgcaagcct cagcgaccga atatatcggt tatgcgtggg cgatggttgt 120
tgtcattgtc ggcgcaacta tcggtatcaa gctgtttaag aaattcacct cgaaagcaag 180
ctgataaacc gatagaattc aagg 204
```

<210> 2

<211> 286

<212> DNA

<211> unknown

<400> 2

```
attggcgcgc tggcaacgat tgcccgtgcg caaggcggcg taatgcgtca tgtcaaaccg 60
cacggcatgt tgtacaacca ggcggcgaaa gaagcacaac tggcagacgc catcgccaga 120
gcggtatacg cttgcgatcc agcattgatt ctcgtcgggc tggcgggaag cgagctgatt 180
cgtgcaggca agcaatatgg tctgacaacg cgcgaggaag tgtttgccga tcgcggttat 240
caggctgacg gctcgctggt gccgcgaagc cagtcaggcg cgttga 286
```

<210> 3

<211> 12

<212> DNA

<211> unknown

<400> 3

cagtcaggac gt 12

<210> 4

<211> 12

<212> DNA

<211> unknown

<400> 4

agaacgcgcc tg 12

<210> 5

<211> 12

<212> DNA

<211> unknown

<400> 5

cgtcgctatt aa 12

<210> 6

<211> 12

<212> DNA

<211> unknown

<400> 6

cagggcgcgt ac 12

<210> 7

<211> 12

<212> DNA

<211> unknown

<400> 7

aaaaaaaaaa ad 12

<210> 8

<211> 12

<212> DNA

Please ensure all <211> responses = match actual number of bases

<211> unknown
<400>8
aaatttaaat tt 12
<210>9
<211>12
<212>DNA
<211> unknown
<400>9
aattaattaa tt

12 ← insert

<210>10
<211>12
<212>DNA
<211> unknown
<400>10
acgacgacga cg 12
<210>11
<211>12
<212>DNA
<211> unknown
<400>11
atatatatat at 12
<210>12
<211>12
<212>DNA
<211> unknown
<400>12
cccccccccc cc 12
<210>13
<211>12
<212>DNA
<211> unknown
<400>13
ccggccggcc gg 12
<210>14
<211>12
<212>DNA
<211> unknown
<400>14
ctctctctct ct 12
<210>15
<211>12
<212>DNA
<211> unknown
<400>15
gggggggggg gg 12
<210>16
<211>12
<212>DNA
<211> unknown
<400>16
gagagagaga ga 12
<210>17
<211>12
<212>DNA
<211> unknown
<400>17
ggccggccgg cc 12

```
<210>18
<211>12
<212>DNA
<211>unknown
<400>18
tttttttttt tt 12
<210>19
<211>12
<212>DNA
<211>unknown
<400>19
tttgggtttg gg 12
<210>20
<211>12
<212>DNA
<211>unknown
<400>20
tgctgctgct gc 12
<210>21
<211>12
<212>DNA
<211>unknown
<400>21
tatatatacc ac 12
<210>22
<211>12
<212>DNA
<211>unknown
<400>22
gggcggcgac ct 12
<210>23
<211>12
<212>DNA
<211>unknown
<400>23
aggtcgccgc cc 12
<210>24
<211>12
<212>DNA
<211>unknown
<400>24
ggggtcgagg gg 12
<210>25
<211>12
<212>DNA
<211>unknown
<400>25
gctaaaadaa aa 12
<210>26
<211>12
<212>DNA
<211>unknown
<400>26
caattctaca ac 12
<210>27
<211>12
<212>DNA
```

<211>unknown
<400>27
acgagcgagc gc 12
<210>28
<211>12
<212>DNA
<211>unknown
<400>28
tataattata at 12
<210>29
<211>12
<212>DNA
<211>unknown
<400>29
attataatta ta 12
<210>30
<211>12
<212>DNA
<211>unknown
<400>30
gatcacctcc tta 13) delete parenthesis
<210>31
<211>12/3
<212>DNA
<211>unknown
<400>31
taaggaggtg atc 13
<210>32
<211>12
<212>DNA
<211>unknown
<400>32
cccacccacc ca 12
<210>33
<211>12
<212>DNA
<211>unknown
<400>33
tgggtgggtg gg 12
<210>34
<211>25
<212>DNA
<211>unknown
<400>34
gaggaaacag ctatgagatc ttctc 25
<210>35
<211>25
<212>DNA
<211>unknown
<400>35
caggaaacag ctatgacgtt ctcac 25
<210>36
<211>16
<212>DNA
<211>unknown
<400>36
ggcgatatcc ctgaaa 16

```
<210>37
<211>16
<212>DNA
<211>unknown
<400>37
tattatttcc gcaaag 16
<210>38
<211>17
<212>DNA
<211>unknown
<400>38
caggaaacag ctatgac 17
<210>39
<211>12
<212>DNA
<211>unknown
<400>39
cy3-agaacgcgcc tg 12
<210>40
<211>18 /17
<212>DNA
<211>unknown
<400>40
FITC-caggaaaca gctatgac 18 /17
<210>41
<211>31
<212>DNA
<211>unknown
<400>41
FITC-tgctacgtct cttccgatgc tgtctttcgc t 31
<210>42
<211>31
<212>DNA
<211>unknown
<400>42
cy3-tgctacgtct cttccgatgc tgtctttcgc t 31
<210>43
<211>12
<212>DNA
<211>unknown
<400>43
HEX-gaacctcccg ac 12
<210>44
<211>12
<212>DNA
<211>unknown
<400>44
TAM-Tgctgctgct gc 12
```

delete

show bases only – no dashes or extraneous material

delete space

use lower-case letters

Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.